

CLAIMS

What is claimed is:

1. A method comprising:

describing one or more camera motions that occurred when sequential frames of a video were captured, the description for each said frame being provided by:

a set of displacement curves that describe the one or more camera motions in respective horizontal (H), vertical (V), and radial (R) directions; and

a mean absolute difference (MAD) curve that describes a minimum MAD value of the set of displacement curves; and

detecting a shot boundary in the sequential frames by an abrupt transition of the minimum MAD values within the sequential frames, wherein the abrupt transition is indicated by a peak in the minimum MAD values described by the MAD curve.

2. A method as described in claim 1, further comprising characterizing each of the described one or more camera motions as one of still, camera shaking, irregular camera motion, and one or more regular camera motions.

3. A method as described in claim 1, wherein the shot boundary is detected by examining a central said frame, denoted by k , of a sliding window containing N of said frames such that when each of the criteria defined by (a), (b), and (c) as follows are true for the minimum MAD value of the central said frame k , the abrupt transition is detected at the central said frame k :

$$(a) \quad MAD(k) = \max(MAD(i)), i = k - N/2, \dots, k + N/2,$$

$$(b) \quad MAD(k) \geq \alpha_{low} MAD_{sm} + \beta ; \text{ and}$$

$$(c) \quad MAD(k) \geq \alpha_{high} \frac{\sum_{i=-N/2, i \neq k}^{N/2} MAD(i)}{N} + \beta ,$$

wherein α_{low} , α_{high} and β are constants, and MAD_{sm} is a second maximum of the N -frames sliding window.

4. A method as described in claim 1, further comprising generating the set of displacement curves by:

comparing the sequential frames, one to another, utilizing an integral matching template to find matching pixels in respective said frames;

determining displacement of the matching pixels in the compared sequential frames; and

describing the displacement utilizing the set of displacement curves.

5. A method as described in claim 1, further comprising segmenting the video into a plurality of segments based on the detected shot boundary, each said segment having one or more said camera motions that are visually consistent.

6. One or more computer-readable media comprising computer-executable instructions that, when executed, perform the method as recited in claim 1.

7. A method comprising:

describing one or more camera motions that occurred when sequential frames of a

video were captured, the description for each said frame being provided by a set of displacement curves that describe the one or more camera motions in respective horizontal (H), vertical (V), and radial (R) directions; and

detecting camera shaking in the video based on the described camera motions in the set of displacement curves.

8. A method as described in claim 7, wherein the detecting includes defining motion energy E and variation Var based on the set of displacement curves.

9. A method as described in claim 7, wherein the detecting includes:
computing motion energy E and variation Var for a segment of the video from the set of displacement curves, wherein the segment includes at least two of the sequential frames;

finding a probabilistic distribution of camera shaking and non-shaking along E and Var axes; and

determining whether the segment includes camera shaking from the probabilistic distribution.

10. A method as described in claim 7, wherein the camera shaking is detected for a segment of the video, denoted as t_1 to t_2 , having at least two of the sequential frames by:

computing motion energy E and variation Var based on the set of displacement curves to find a distribution of shaking and non-shaking in E - Var space, wherein E and

Var are computed as follows:

$$E = \frac{1}{t_2 - t_1} \sqrt{\sum_x \left(\int_{t_1}^{t_2} |x| dt \right)^2}, x \in \{H, V, R\}; \text{ and}$$

$$Var = \frac{1}{t_2 - t_1} \sum_x \left(\int_{t_1}^{t_2} |\nabla x| dt \right), x \in \{H, V, R\}$$

wherein “ ∇ ” is a differential operator to compute relative variation;

finding a probabilistic distribution of camera shaking and non-shaking along E and Var axes by:

quantifying both the E axis and the Var axis to form a plurality of levels;

and

accumulating motion energy E and variation Var in respective said levels to form the probabilistic distribution; and

determining whether the segment includes camera shaking from the probabilistic distribution by:

computing average error probability P_E wherein:

$$P_E = P(S) \int_{R_{\bar{S}}} p(x|S) dx + P(\bar{S}) \int_{R_S} p(x|\bar{S}) dx$$

S is a hypothesis that the segment is shaking;

\bar{S} is a hypothesis that the segment is non-shaking;

$P(S)$ and $P(\bar{S})$ are prior probabilities for existence of shaking and non-shaking, respectively;

$R_{\bar{S}}$ is a missed detection area; and

R_S is a false detection area; and

obtaining optimal decisions T_E and T_V based on a minimization of P_E by a likelihood ratio test wherein:

if the segment satisfies optimal decisions T_E and T_V , the segment includes camera shaking; and

the optimal decisions T_E and T_V obtained as follows:

$$T = \log \frac{P(S)P(E|S)}{P(\bar{S})P(E|\bar{S})} \begin{matrix} \bar{S} \\ < \\ S \end{matrix} 0.$$

11. A method as described in claim 7, further comprising generating the set of displacement curves by:

comparing the sequential frames, one to another, utilizing an integral matching template to find matching pixels in respective said frames;

determining displacement of the matching pixels in the compared sequential frames; and

describing the displacement utilizing the set of displacement curves.

12. One or more computer-readable media comprising computer-executable instructions that, when executed, perform the method as recited in claim 7.

13. A method comprising:

describing one or more camera motions that occurred when sequential frames of a video were captured, the description for each said frame being provided by:

a set of displacement curves that describe the one or more camera motions in respective horizontal (H), vertical (V), and radial (R) directions;

a mean absolute difference (MAD) curve that describes a minimum MAD value of the set of displacement curves; and

a MAJ curve that is a qualitative description of the one or more camera motions and is determined from the minimum MAD value of the MAD curve; and characterizing each of the described one or more camera motions as one of still, camera shaking, irregular camera motion, and one or more regular camera motions.

14. A method as described in claim 13, wherein the qualitative description of camera motion includes still, horizontal, vertical and radial.

15. A method as described in claim 13, further comprising generating the set of displacement curves by:

comparing the sequential frames, one to another, utilizing an integral matching template to find matching pixels in respective said frames;

determining displacement of the matching pixels in the compared sequential frames; and

describing the displacement utilizing the set of displacement curves.

16. A method as described in claim 13, wherein the characterizing includes: morphological filtering of the set of displacement curves by:

performing an opening operation to remove unstable motions from the set of displacement curves to form a set of regular motion displacement curves; and

performing a closing operation on the set of regular motion displacement

curves to connect interrupted regular motions to form a set of connected regular motion displacement curves;

detecting shot boundaries based on the set of connected regular motion displacement curves; and

merging overlap segments of the set of connected regular motion displacement curves to form one or more motion segments which describe regular camera motions that are included in at least two of the set of connected regular motion displacement curves.

17. A method as described in claim 13, wherein the characterizing further comprises:

defining motion energy E and variation V based on the set of displacement curves to detect the camera shaking and the irregular camera motion; and

detecting the still camera motion from the MAJ curve.

18. A method as described in claim 13, wherein the one or more regular camera motions are selected from the group consisting of:

panning;

tilting;

zooming;

tracking;

booming;

dollying; and

any combination thereof.

19. One or more computer-readable media comprising computer-executable instructions that, when executed, perform the method as recited in claim 13.

20. A method comprising:

analyzing video having sequential frames to determine one or more camera motions that occurred when sequential frames of the video were captured; and

describing the one or more camera motions for each said frame by:

- a set of displacement curves that describe the one or more camera motions in respective horizontal (H), vertical (V), and radial (R) directions;
- a mean absolute difference (MAD) curve that relates a minimum MAD value from the set of displacement curves; and
- a major motion (MAJ) curve that:
 - is generated from the minimum MAD value; and
 - provides one or more qualitative descriptions that describe the one or more camera motions as one of still, vertical, horizontal and radial.

21. A method as described in claim 20, further comprising detecting a shot boundary in the sequential frames from an abrupt transition of respective said minimum MAD values.

22. A method as described in claim 20, further comprising detecting a shot boundary in the sequential frames from an abrupt transition by examining a central said

frame, denoted by k , of a sliding window containing N of said frames such that when each of the criteria defined by (a), (b), and (c) as follows are true for the minimum MAD value of the central said frame k , the abrupt transition is detected at the central said frame k :

$$(a) \quad MAD(k) = \max(MAD(i)), i = k - N/2, \dots, k + N/2;$$

$$(b) \quad MAD(k) \geq \alpha_{low} MAD_{sm} + \beta; \text{ and}$$

$$(c) \quad MAD(k) \geq \alpha_{high} \frac{\sum_{i=-N/2, i \neq k}^{N/2} MAD(i)}{N} + \beta,$$

wherein α_{low} , α_{high} and β are constants, and MAD_{sm} is a second maximum of the N -frames sliding window.

23. A method as described in claim 20, further comprising detecting camera shaking in the video by defining motion energy E and variation Var based on the set of displacement curves.

24. A method as described in claim 20, further comprising detecting camera shaking by:

computing motion energy E and variation Var for a segment of the video from the set of displacement curves, wherein the segment includes at least two of the sequential frames;

finding a probabilistic distribution of camera shaking and non-shaking along E and Var axes; and

determining whether the segment includes camera shaking from the probabilistic distribution.

25. A method as described in claim 20, further comprising detecting camera shaking for a segment of the video, denoted as t_1 to t_2 , having at least two of the sequential frames by:

computing motion energy E and variation Var based on the set of displacement curves to find a distribution of shaking and non-shaking in E - Var space, wherein E and Var are computed as follows:

$$E = \frac{1}{t_2 - t_1} \sqrt{\sum_x \left(\int_{t_1}^{t_2} |x| dt \right)^2}, x \in \{H, V, R\}; \text{ and}$$

$$Var = \frac{1}{t_2 - t_1} \sum_x \left(\int_{t_1}^{t_2} |\nabla x| dt \right), x \in \{H, V, R\}$$

wherein “ ∇ ” is a differential operator to compute relative variation;

finding a probabilistic distribution of camera shaking and non-shaking along E and Var axes by:

quantifying both the E axis and the Var axis to form a plurality of levels;

and

accumulating motion energy E and variation Var in respective said levels to form the probabilistic distribution; and

determining whether the segment includes camera shaking from the probabilistic distribution by:

computing average error probability P_E wherein:

$$P_E = P(S) \int_{R_{\bar{S}}} p(x|S) dx + P(\bar{S}) \int_{R_S} p(x|\bar{S}) dx$$

S is a hypothesis that the segment is shaking;

\bar{S} is a hypothesis that the segment is non-shaking;

$P(S)$ and $P(\bar{S})$ are prior probabilities for existence of shaking and non-shaking, respectively;

$R_{\bar{S}}$ is a missed detection area; and

R_S is a false detection area; and

obtaining optimal decisions T_E and T_V based on a minimization of P_E by a likelihood ratio test wherein:

if the segment satisfies optimal decisions T_E and T_V , the segment includes camera shaking; and

the optimal decisions T_E and T_V obtained as follows:

$$T = \log \frac{P(S)P(E|S)}{P(\bar{S})P(E|\bar{S})} \begin{matrix} \bar{S} \\ < \\ S \end{matrix} 0 .$$

26. A method as described in claim 20, further comprising characterizing each of the described one or more camera motions as one of still camera motion, camera shaking, irregular camera motion, and one or more regular camera motions.

27. A method as described in claim 26, wherein the one or more regular camera motions are selected from the group consisting of:

panning;

tilting;

zooming;
tracking;
booming;
dollying; and
a combination thereof.

28. A method as described in claim 20, further comprising characterizing each of the described one or more camera motions as one or more regular camera motions by:

morphological filtering of the set of displacement curves by:

performing an opening operation to remove unstable motions from the set of displacement curves to form a set of regular motion displacement curves; and

performing a closing operation on the set of regular motion displacement curves to connect interrupted regular motions to form a set of connected regular motion displacement curves;

detecting shot boundaries based on the set of connected regular motion displacement curves; and

merging overlap segments of the set of connected regular motion displacement curves to form one or more motion segments which describe regular camera motions that are included in at least two of the set of connected regular motion displacement curves.

29. A method as described in claim 20, wherein the analyzing is performed by:

executing a *Mean Absolute Difference* (MAD) algorithm based on luminance (L),
wherein:

luminance (L) is utilized to match pixels in respective first and second said
frames; and

the MAD between the first said frame, denoted as k , and the second said
frame, denoted as $k+1$, is computed as follows:

$$MAD(\Delta x) = \frac{1}{N} \sum_{x \in T} |L(x + \Delta x, k + 1) - L(x, k)|$$

wherein T is a set of the matching pixels and “ N ” is a total number of the
matching pixels; and

estimating camera motion through error minimization in parameter state space (H ,
 V , R), which respectively denote the horizontal, the vertical, and the radial directions,
wherein:

for each said direction (H , V , R), optimal matching is obtained when a
minimum MAD is reached as follows:

$$(H, V, R) = \underset{\Delta x \in \Psi}{\operatorname{argmin}} MAD(\Delta x)$$

Ψ is a state space of a plurality of camera motion candidates; and

the camera motion candidate that yields a best match is an optimal
estimation of camera motion.

30. A method as described in claim 20, further comprising generating the set
of displacement curves by:

comparing the sequential frames, one to another, utilizing an integral matching

template to find matching pixels in respective said frames;

determining displacement of the matching pixels in the compared sequential frames; and

describing the displacement utilizing the set of displacement curves.

31. One or more computer-readable media comprising computer-executable instructions that, when executed, perform the method as recited in claim 20.

32. A computer comprising:

video having sequential frames;

a set of displacement curves that describe one or more camera motions that occurred when sequential frames of a video were captured in respective horizontal (H), vertical (V), and radial (R) directions;

a mean absolute difference (MAD) curve that relates a minimum MAD value of the set of displacement curves for each said frame; and

a major motion (MAJ) curve that is a qualitative description of camera motion for each said frame and is determined from the minimum MAD value of the MAD curve; and

a video analysis module this is executable to:

detect a shot boundary in the sequential frames from an abrupt transition of respective said minimum MAD values;

detect camera shaking in the sequential frames based on the described camera motions in the set of displacement curves; and

characterize each of the one or more camera motions as one of still camera motion, camera shaking, irregular camera motion, and one or more regular camera motions.

33. A computer as described in claim 32, wherein the video analysis module is executable to generate the set of displacement curves by:

comparing the sequential frames, one to another, utilizing an integral matching template to find matching pixels in respective said frames;

determining displacement of the matching pixels in the compared sequential frames; and

describing the displacement utilizing the set of displacement curves.

34. A computer as described in claim 32, wherein the video analysis module is executable to:

executing a *Mean Absolute Difference* (MAD) algorithm based on luminance (L), wherein:

luminance (L) is utilized to match pixels in respective first and second said frames; and

the MAD between the first said frame, denoted as k , and the second said frame, denoted as $k+1$, is computed as follows:

$$MAD(\Delta x) = \frac{1}{N} \sum_{x \in T} |L(x + \Delta x, k + 1) - L(x, k)|$$

wherein T is a set of the matching pixels and “ N ” is a total number of the matching pixels; and

estimating camera motion through error minimization in parameter state space (H, V, R) , which respectively denote the horizontal, the vertical, and the radial directions, wherein:

for each said direction (H, V, R) , optimal matching is obtained when a minimum MAD is reached as follows:

$$(H, V, R) = \underset{\Delta x \in \Psi}{\operatorname{argmin}} \operatorname{MAD}(\Delta x)$$

Ψ is a state space of a plurality of camera motion candidates; and

the camera motion candidate that yields a best match is an optimal estimation of camera motion.

35. A computer as described in claim 32, wherein the video analysis module is executable to detect the shot boundary by examining a central said frame, denoted by k , of a sliding window containing N of said frames such that when each of the criteria defined by (a), (b), and (c) as follows are true for the minimum MAD value of the central said frame k , the abrupt transition is detected at the central said frame k :

$$(a) \quad \operatorname{MAD}(k) = \max(\operatorname{MAD}(i)), i = k - N/2, \dots, k + N/2;$$

$$(b) \quad \operatorname{MAD}(k) \geq \alpha_{low} \operatorname{MAD}_{sm} + \beta; \text{ and}$$

$$(c) \quad \operatorname{MAD}(k) \geq \alpha_{high} \frac{\sum_{i=-N/2, i \neq k}^{N/2} \operatorname{MAD}(i)}{N} + \beta,$$

wherein α_{low} , α_{high} and β are constants, and MAD_{sm} is a second maximum of the N -frames sliding window.

36. A computer as described in claim 32, wherein the video analysis module is

executable to detect camera shaking by defining motion energy E and variation Var based on the set of displacement curves.

37. A computer as described in claim 32, wherein the video analysis module is executable to detect camera shaking by:

computing motion energy E and variation Var for a segment of the video from the set of displacement curves, wherein the segment includes at least two of the sequential frames;

finding a probabilistic distribution of camera shaking and non-shaking along E and Var axes; and

determining whether the segment includes camera shaking from the probabilistic distribution.

38. A computer as described in claim 32, wherein the video analysis module is executable to detect camera shaking for a segment of the video, denoted as t_1 to t_2 , having at least two of the sequential frames by:

computing motion energy E and variation Var based on and the set of displacement curves to find a distribution of shaking and non-shaking in E - Var space, wherein E and Var are computed as follows:

$$E = \frac{1}{t_2 - t_1} \sqrt{\sum_x \left(\int_{t_1}^{t_2} |x| dt \right)^2}, x \in \{H, V, R\}; \text{ and}$$
$$Var = \frac{1}{t_2 - t_1} \sum_x \left(\int_{t_1}^{t_2} |\nabla x| dt \right), x \in \{H, V, R\}$$

wherein “ ∇ ” is a differential operator to compute relative variation;

finding a probabilistic distribution of camera shaking and non-shaking along E and Var axes by:

quantifying both the E axis and the Var axis to form a plurality of levels;

and

accumulating motion energy E and variation Var in respective said levels to form the probabilistic distribution; and

determining whether the segment includes camera shaking from the probabilistic distribution by:

computing average error probability P_E wherein:

$$P_E = P(S) \int_{R_{\bar{S}}} p(x|S) dx + P(\bar{S}) \int_{R_S} p(x|\bar{S}) dx$$

S is a hypothesis that the segment is shaking;

\bar{S} is a hypothesis that the segment is non-shaking;

$P(S)$ and $P(\bar{S})$ are prior probabilities for existence of shaking and non-shaking, respectively;

$R_{\bar{S}}$ is a missed detection area; and

R_S is a false detection area; and

obtaining optimal decisions T_E and T_V based on a minimization of P_E by a likelihood ratio test wherein:

if the segment satisfies optimal decisions T_E and T_V , the segment includes camera shaking; and

the optimal decisions T_E and T_V obtained as follows:

$$T = \log \frac{P(S)P(E|S)}{P(\bar{S})P(E|\bar{S})} \begin{matrix} \bar{S} \\ < \\ > \\ S \end{matrix} 0 .$$

39. A computer as described in claim 32, wherein the video analysis module is executable to characterize by:

morphological filtering of the set of displacement curves by:

performing an opening operation to remove unstable motions from the set of displacement curves to form a set of regular motion displacement curves; and

performing a closing operation on the set of regular motion displacement curves to connect interrupted regular motions to form a set of connected regular motion displacement curves;

detecting shot boundaries based on the set of connected regular motion displacement curves; and

merging overlap segments of the set of connected regular motion displacement curves to form one or more motion segments which describe regular camera motions that are included in at least two of the set of connected regular motion displacement curves.

40. A computer as described in claim 32, wherein the regular camera motions are selected from the group consisting of:

panning;

tilting;

zooming;

tracking;

booming;

dolly; and

a combination thereof.

41. A computer as described in claim 32, wherein the video analysis module is executable to characterize by:

defining motion energy E and variation V based on the set of displacement curves to detect the camera shaking and the irregular camera motion; and

detecting the still camera motion from the MAJ curve.